

ОСОБЛИВОСТІ АРХІТЕКТУРИ DEEPLABV3+

Прочухан Д.В.

Національний технічний університет

«Харківський політехнічний інститут», м. Харків

DeepLabv3+ - архітектура згорткової нейронної мережі, призначена для сегментації зображень. Вказана модель синтезує переваги архітектур DeepLabv3 та U-Net, забезпечуючи покращену точність сегментації.

Архітектура DeepLabv3+ успадковує модуль Atrous Spatial Pyramid Pooling від моделі DeepLabv3. Atrous Spatial Pyramid Pooling дозволяє мережі захоплювати контекстну інформацію різного масштабу за допомогою згортки. Atrous Spatial Pyramid Pooling складається з паралельних гілок, кожна з яких використовує згортку з різними розмірами ядра. Наведена згортка вставляє пропуски між елементами фільтра, що дозволяє їй охоплювати ширше поле зору без збільшення розміру ядра. Вказана особливість Atrous Spatial Pyramid Pooling надає можливість збирати інформацію з різних масштабів, що покращує розуміння моделлю елементів зображення. На наступному кроці виходи з паралельних гілок об'єднуються за допомогою каскадної згортки. Таким чином, модуль Atrous Spatial Pyramid Pooling є одним з ключових компонентів архітектури DeepLabv3+, який дозволяє краще сегментувати об'єкти з різними розмірами та формами. Модель DeepLabv3+ використовує архітектуру кодер-декодер. Кодер відповідає за вилучення ознак з зображення. Кодер складається з послідовних згорткових та пулінгових шарів, які поступово зменшують просторову роздільну здатність зображення, але збільшують його глибину представлення. Декодер використовує ознаки для створення сегментаційної карти. Декодер складається з послідовних шарів деконволюції та з'єднання, які поступово збільшують просторову роздільну здатність сегментаційної карти. Таким чином, модель DeepLabv3+ поєднує ознаки з енкодера з ознаками з декодера. Вказані особливості дозволяють моделі DeepLabv3+ одночасно вивчати глобальні та локальні ознаки зображення, що призводить до більш точної сегментації. Переваги архітектури DeepLabv3+ наступні: висока точність сегментації для широкого спектра задач, здатність обробляти зображення з різними розмірами та формами, можливість використання різних базових мереж, ефективність та відносно невелика обчислювальна складність. Недоліки моделі DeepLabv3+ наступні: чутливість до шуму в даних, вимоги до великої кількості даних для навчання. Архітектура DeepLabv3 має значні можливості для розв'язку широкого кола задач сегментації.

Література:

1. YURTKULU, Salih Can; ŞAHIN, Yusuf Hüseyin; UNAL, Gozde. Semantic segmentation with extended DeepLabv3 architecture. In: 2019 27th Signal Processing and Communications Applications Conference (SIU). IEEE, 2019. p. 1-4.